Revision 7.0

MT&R Guidelines

Monitoring, Targeting and Reporting (MT&R) Reference Guide

Prepared by: ESI Energy Performance Tracking (EPT) Team

October 19, 2018

Acknowledgements

The authors would like to thank the following individuals for their contributions to this document:

Todd Amundson	Bonneville Power Administration
Steve Brooks	Peninsula Light Co. (formerly Bonneville Power Administration)
Jennifer Langdon	Cowlitz PUD (formerly Cascade Energy, Inc.)
Keri Macklin	CLEAResult (formerly Triple Point Energy, Inc.)
Steve Martin	Cascade Energy, Inc.
Steve Mulqueen	Cascade Energy, Inc.
Jacob Schroeder	Cascade Energy, Inc.
Sara York	Cascade Energy, Inc.

Contents

About The EPT Team

The Energy Performance Tracking (EPT) team is responsible for defining and documenting the Monitoring, Targeting, and Reporting (MT&R) methodologies employed in the Energy Smart Industrial's (ESI) Strategic Energy Management (SEM) project implementation and maintaining the contents of this document. The EPT team is chaired by the Bonneville Power Administration (BPA) Energy Management Engineer and includes participants from BPA's Energy Efficiency team and implementation Program Partner(s).

Introduction

The Energy Smart Industrial (ESI) program uses a Monitoring, Targeting, and Reporting (MT&R) methodology—in conjunction with a process to track specific activities—to estimate energy savings for Strategic Energy Management (SEM) projects. This document outlines recommended methodologies to 1) establish baseline energy models at a whole-facility or subsystem level, and 2) quantify energy savings associated with the implementation of multiple energy efficiency measures (EEMs) over a defined performance period.

In the context of ESI whole-facility or subsystem energy management, the default approach is a top-down, forecasting-based regression model as described by the International Performance Measurement and Verification Protocol (IPMVP).[1] Unless otherwise noted, the ESI MT&R Guidelines are intended to align with the best practices outlined by IPMVP for "Option C" models.

Developing a linear regression model to monitor and report energy savings for industrial SEM projects while maintaining consistency with IPMVP is an iterative process. This process requires the practitioner to work with large data sets, to understand the major energy drivers in a facility, and to have a working knowledge of statistics. The predictive ability of the model depends largely upon the practitioner's ability to navigate this iterative process in a sequential manner.

Sections 1-3 of this document focus on the model development process. Sections 4-6 of this document focus on the quantification of energy savings attributable to SEM. Specific focus is given to addressing the separation of operations and maintenance savings from concurrent capital projects and adjusting the baseline model for non-routine changes to plants or systems.

1 *International Performance Measurement and Verification Protocol.* Efficiency Evaluation Organization. 10000-1:2016. www.evo-world.org.

1. Characterizing the Facility or Process

1.1 Identify Measurement Boundary

- For whole-facility energy models, the measurement boundary consists of all the systems and processes served by one or more utility meters. While energy sources may include natural gas, steam, or compressed air, the examples in this document assume electrical energy as the targeted response variable.
- Care must be taken to ensure that:
    - All electrical energy crossing the measurement boundary has been documented and accounted for. Documentation may include one-line electrical drawings, energy maps, and system schematics which identify equipment and processes within the measurement boundary.
    - Significant electrical energy-consuming equipment within the measurement boundary that inconsistently supplies other areas of the plant is documented and accounted for. An example is an air compressor within the measurement boundary that supplies variable amounts of compressed air to equipment both within the measurement boundary and other areas of the plant. In such cases, effective sub-metering strategies need to be deployed to measure the energy usage crossing the measurement boundary for reporting purposes.
    - If other energy sources are used to offset electrical energy use within the measurement boundary, then effective sub-metering strategies must be deployed to measure the changing energy sources for reporting purposes. One such example is a drying process that can use a fan, a steam heater, or a combination of both.

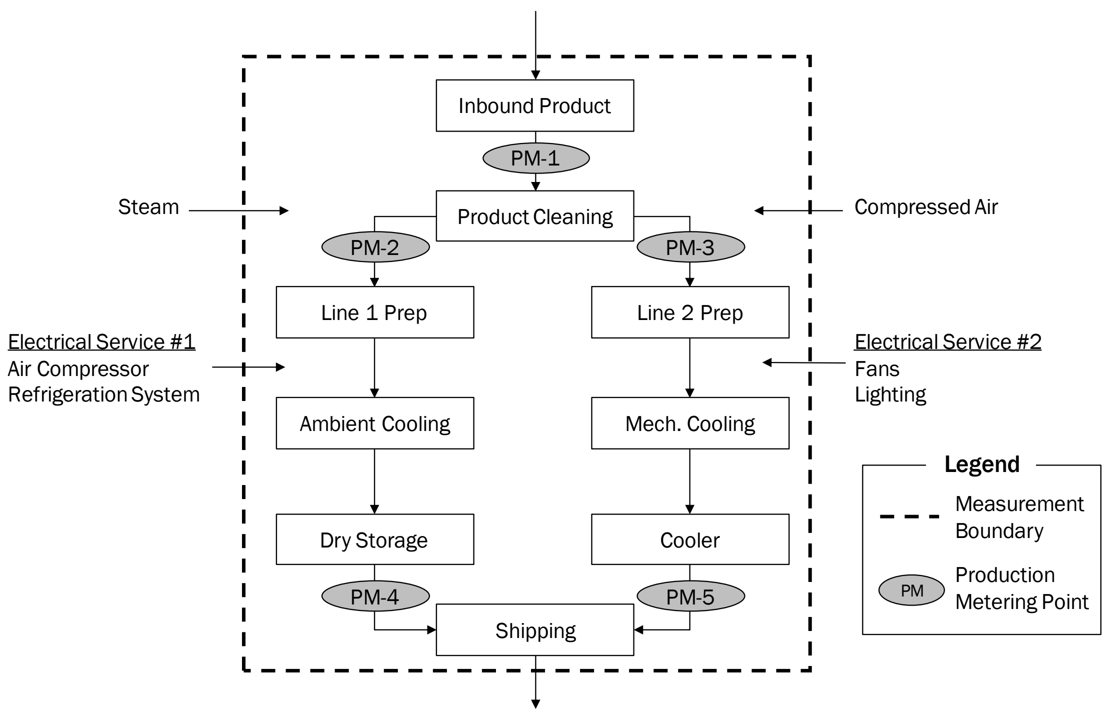


Figure 1. Illustration of measurement boundary, including where product, energy, steam, and compressed air cross the measurement boundary.

1.2 Identify Production Energy Drivers

- The primary energy driver is typically production. It is important to quantify how many product types are manufactured in the facility and understand whether there is likely to be a difference in energy intensity based on lead time, process flow, batch size, etc. Raw material, work in progress, and finished product metrics each have merits and demerits for selection as primary energy driver variables. An informed decision will consider factors such as lead time, the desire to account for yield effects, and the prevalence of inventory fluctuations in-process or at the finished product stage.
- The source of production data must be understood to assess how it physically relates to the energy intensive processes. If a significant offset exists between the energy-intensive process step and the production measurement gate, a compensating time-series shift that corresponds to the magnitude of the time offset may be applied (see Section 2.3).
- Process flow diagrams, piping and instrumentation diagrams, and value stream maps can be helpful at this stage.

Table 1. Consideration for Selection of Production Variable

MEASUREMENT GATE	MERIT	DEMERIT
Raw material input	Provides a mechanism to capture the effects of different raw material types.	Will not produce a signal for energy impact of yield or productivity improvements.
Work in progress	Allows selection of production variable at energy-intensive process step, thereby minimizing time series shift.	Availability of data may be limited. Does not provide mechanism for incentivizing energy impact of yield/productivity improvement downstream from point of measurement.
End of line metric	Provides mechanism for incentivizing energy impact of yield/productivity improvements.	May induce a time-series shift for long lead-time processes.
Finished product shipped	Reliable data is typically available from business systems.	May not correspond with production if finished product inventory fluctuates.

1.3 Identify Other Energy Drivers – Hypothesis Stage

- Based on the system inventory and process characteristics, form a hypothesis of other energy drivers. The most common examples are ambient conditions (dry-bulb and wet-bulb temperatures) but can include variables such as raw material properties, operational modes (weekend/day), occupancy, etc.
- Energy drivers must be tested for statistical significance (see Section 3.1). A suitable explanation must be provided if an energy driver that is not statistically significant is nevertheless used in the model.
- Ambient temperature must be tested for statistical significance. If temperature is omitted from the model, the rationale must be documented.
- In the process of variable selection, the model developer will face competing objectives: capture the full subset of statistically significant variables and provide the customer with a model that is simple and easy to maintain. No single selection criteria will provide the perfect solution, so the modeler must rely on his or her experience and engineering judgment.
- Including process variables in the energy model may add to the explanatory power of the model but can limit the ability to measure savings. If a process variable is included in the model and a key EEM has a direct impact on this variable, then the energy savings measured using this model are likely to

be inaccurate. While sometimes necessary for model fitness, including process variables is not a preferred option.

- See Figure 2 for example. Blast freezers generally operate at less than full capacity, and runtime may trend with energy consumption. An energy efficiency measure exists to reduce freezer run time. However, if the number of run hours is included as a model variable, the savings from this opportunity would not be estimated. Pounds of product frozen would be a more appropriate variable to include.

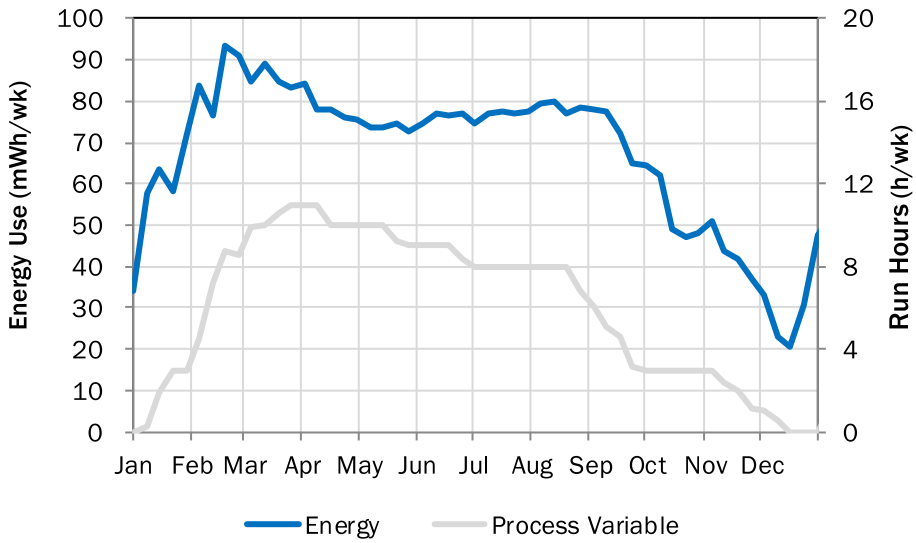


Figure 2. Example of energy use and process variable tracking. Like energy use, the process variable can be influenced by energy efficiency measures.

1.3.1 Weather Data

Acceptable sources of weather data include the National Climate Data Center (NCDC) and Weather Underground. Use of weather data from Energy Management and Information Systems (EMIS) that agree with these sources is also acceptable. A change in the weather data source during the reporting period should trigger an update to the original model, followed by EPT team review.

1.4 Identify Utility Meters or Submeters

- Document which processes are served by specific meters. This step will be important in determining whether to create a single model for a facility or to create discrete models for functional units that collectively represent the entire facility's energy use.
- Meter serial numbers, utility account numbers, or other unique identifiers must be recorded in the baseline report.
- If an end user-owned submeter will be used in place of the utility meter, the submeter data should be appropriately aggregated and compared to a utility bill. If the submetered measurement boundary does not align with a utility meter, then meter calibration should be confirmed by a certified electrician. The electrician shall strive to use no less than third-order, NIST-traceable calibration equipment, as recommended by ASHRAE Guideline 14-2014, Section 6.4.2.[2]

[2] *ASHRAE Guideline 14–Measurement of Energy, Demand, and Water Savings.* American Society of Heating, Refrigerating and Air Conditioning Engineers. 2014.

2. Establishing a Baseline Data Set

2.1 Determine the Baseline Period

- The baseline period should encompass the cycles and ranges of the hypothesized primary and secondary energy drivers and extend as close to the start of the reporting period as possible. Ideally, the baseline period captures two or more cycles of operation.
- If re-baselining is required for participants re-enrolling in SEM, the last reporting period of the previous engagement is typically used for the new baseline period.
- The guideline for the minimum number of baseline data points is: 6 × number of coefficients in the model. If the data set falls below this guideline, the model will likely be "over-fitted", and the model's comparative performance will likely deteriorate during the reporting period. Since the number of coefficients is not known at this point, it can be assumed that there will be one coefficient for each hypothesized variable, plus the intercept.
- Energy use that exhibits seasonal dependence should use complete years (e.g. 12, 24, or 36 months) of continuous data during the baseline period to ensure balanced representation of all operating modes. Models that use other ranges of baseline data can create statistical bias by under- or over-representing normal modes of operation.[3]
- Data with daily or weekly time resolutions typically provide better insights about processes, and thus result in more accurate models when compared to data of longer durations such as monthly data. Process lead time should be considered when selecting the modeling interval, both for determining the modeling interval and applying time-series offsets with the corresponding energy data.
- The NW Strategic Energy Management Collaborative white paper, "Common Considerations in Defining Baselines for Industrial Strategic Energy Management Projects," provides additional guidance and case studies on the selection of an appropriate baseline period and the treatment of non-production periods in a daily model.[4]

2.1.1 Temporary and Permanent Baseline Events – Addressing Non-Routine Events, Incentivized or Non-Incentivized Energy Projects

Utility records should be reviewed to confirm whether incentivized energy projects occurred within the measurement boundary during the proposed baseline period. If so, project records should be obtained to accurately capture implementation dates and magnitude of verified savings.

To determine the effective date for an incentivized EEM, apply the earlier of the project measurement and verification (M&V) start date, or the date that an inflection is observed in the energy data (see Appendix A).

Interviews with the end user and serving utility should be conducted to determine if other non-incentivized energy projects occurred during the proposed baseline period. If either case is identified, one of the options in Appendix A can be applied to ensure savings are not double counted.

3 *International Performance Measurement and Verification Protocol.* Efficiency Evaluation Organization. 10000-1:2012. Section 4.8.4.
4 *Common Considerations in Defining Baselines for Industrial Strategic Energy Management Projects.* NW Industrial Strategic Energy Management (SEM) Collaborative. 2014.

2.2 Collect and Review Data

- When collecting data for energy or energy drivers, ensure that accurate records are maintained regarding the data source (e.g., end user database, production gate, weather station).
- Perform an initial review for outliers by plotting each variable independently in a time series format. Identify and flag erroneous entries. Control limits of three standard deviations, ±3 sigma (σ), from the mean are often useful for identifying outliers in normally distributed data.

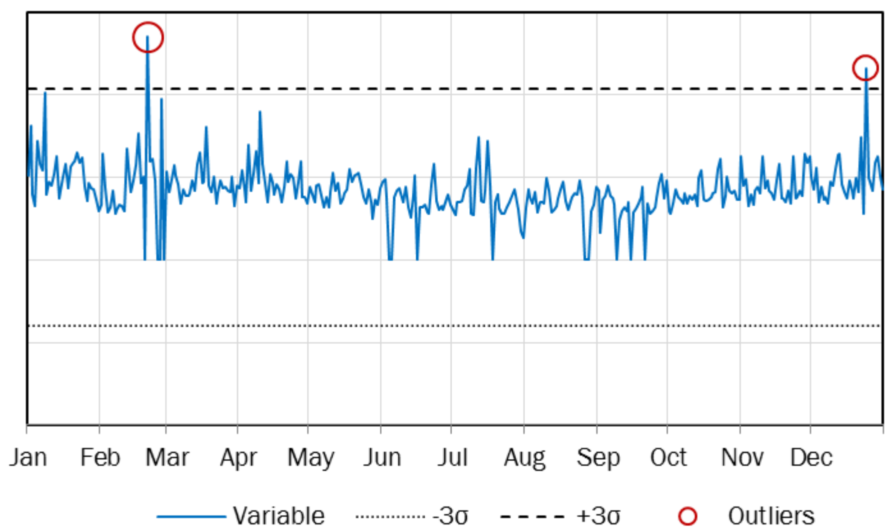


Figure 3. Example of graphical method to identify anomalies.

- Missing data points or data entry errors should be investigated and corrected by the facility, if possible.
- Any outliers that are ultimately removed from the baseline data set should be annotated with assignable cause. Understanding assignable cause will likely require communication with the end user's Energy or Data Champion.
- Generally, avoid replacing missing or outlier data with estimated values. Exceptions are permissible when data is provided at a much finer interval than the model e.g., if time interval of data is 15 minutes or hourly. For energy data, best practice is that values in aggregate match a known reference such as utility billing history.
- Examine data obtained from industrial control systems with a higher level of scrutiny. This data is often hourly or sub-hourly and frequently includes the following types of "bad data":
    - Erroneous values: a value such as "Control System Error"
    - Null values: no data for the given variable and observation
    - Anomalous operations: values that appear out of range of normal operations. This may include values that remain constant when equipment is off.
- Observations that appear anomalous should be reviewed with plant personnel to better understand the operation of the system.
- If any data point within the observation is deemed erroneous, null, or anomalous, the observation should be removed and documented in the Energy Model Report. If the number of observations per time period vary due to removal of invalid data, a weighted regression can be considered as outlined in Appendix E.

- Graphing data can be an effective way to detect erroneous and anomalous data. For example, in Figure 4, power within the dashed box is considerably lower than power above the dashed box for similar machine speeds. This suggests that the operation of this machine should be investigated prior to performing calculations.

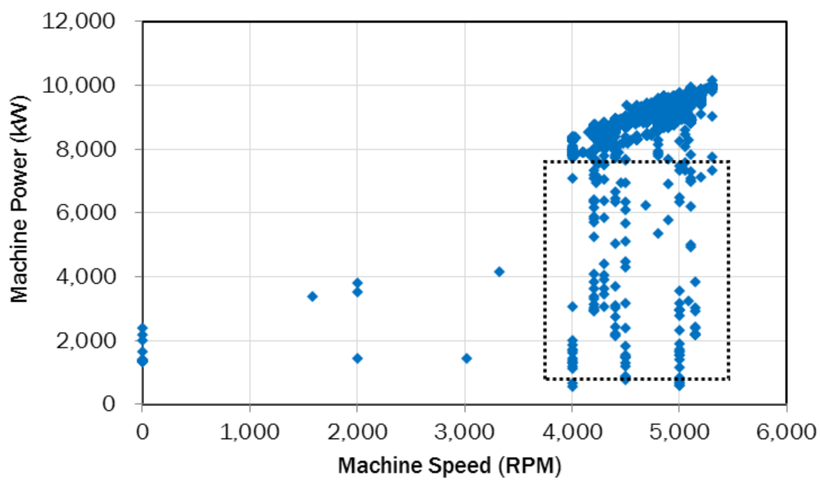


Figure 4. Illustration of control system data showing machine power vs. machine speed.

2.3 Adjust for Time-Series Offsets

- Use time-series plots to identify consistent offsets between energy use and independent variables. For example, if the energy-intensive process is two days' lead time from the production measurement point, a two-day time series adjustment may need to be applied to the production variable. However, this approach may be unnecessary if a longer model interval is selected (e.g., weekly versus daily model).

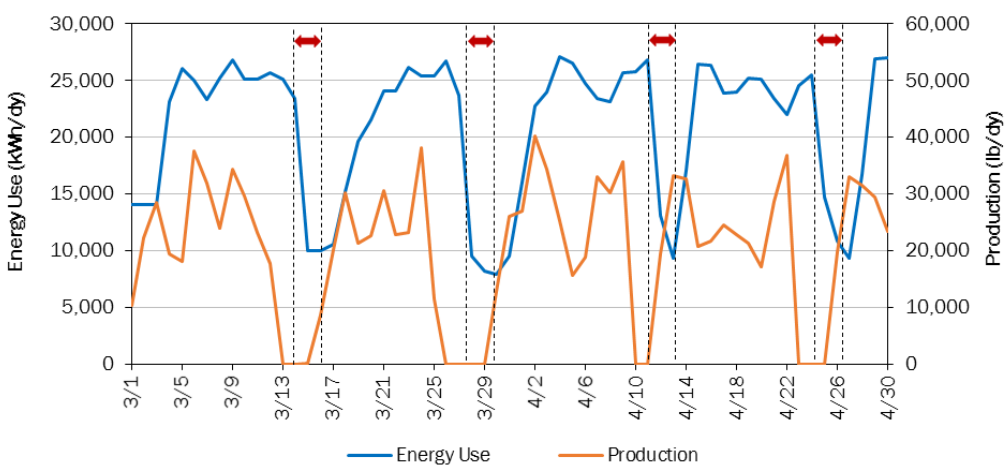


Figure 5. Example of a time-series off-set (energy and production vs. time).

- If necessary, apply the time-series offset to the relevant independent variable(s), maintaining the original source data in a separate file.
- At this point, the baseline data set is ready for the regression modeling process.

2.4 Form a Hypothesis Model

- The hypothesis model should be driven by an informed understanding of the physical characteristics of the process.
- Non-linear and interactive terms should be evaluated when suggested by the data.
- Use scatter diagrams to understand the relationship between energy use and energy drivers. For example, a plant's energy intensity often becomes progressively more efficient at higher production volumes. This implies a non-linear relationship between energy use and production and is illustrated in Figure 6.

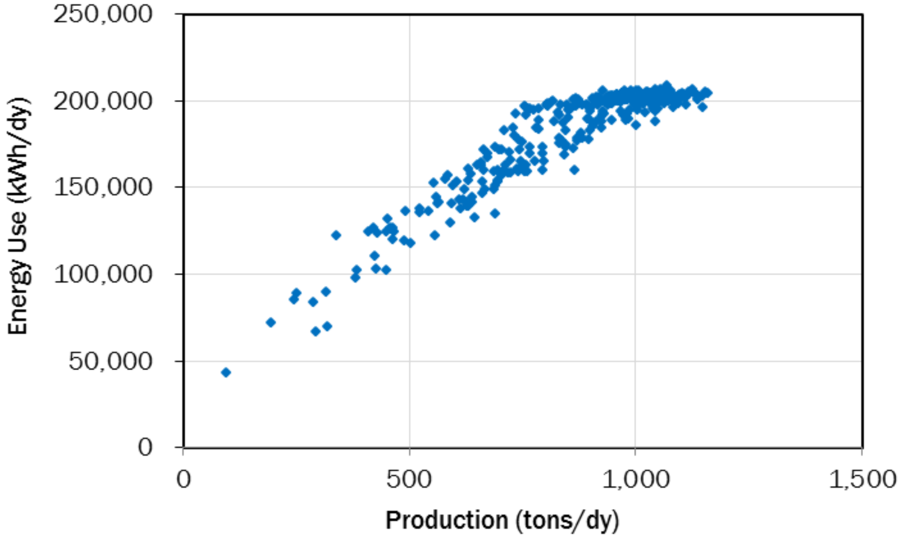


Figure 6. Example of a scatter plot (energy vs. production).

- The energy profiles of facilities with large space conditioning and refrigeration loads often exhibit a "change-point" characteristic. Modeling a facility that exhibits a change-point with a single linear model would introduce unnecessary error. Instead, this system should be modeled with a change-point model. The presence of a change-point can be identified by plotting energy use versus ambient temperature, as illustrated in Figure 7.
- For models with daily time resolution, there is no loss in information when using a change-point model over a degree-day model. For longer time periods, the differences between the two approaches are generally slight, except in mild climates with many temperatures near the balance-point.[5] Therefore, consider a degree-day approach when energy use is driven by temperature and the facility is in a mild climate.
- When two or more independent variables exhibit correlation, multicollinearity is present within the model. The presence of collinear variables can affect the precision of individual coefficients and can understate the statistical significance of individual predicator variables.

5 *Discussion Regarding the Use of Average Temperature or Degree-Days in Energy Regressions.* SBW Consulting. November 28, 2015.

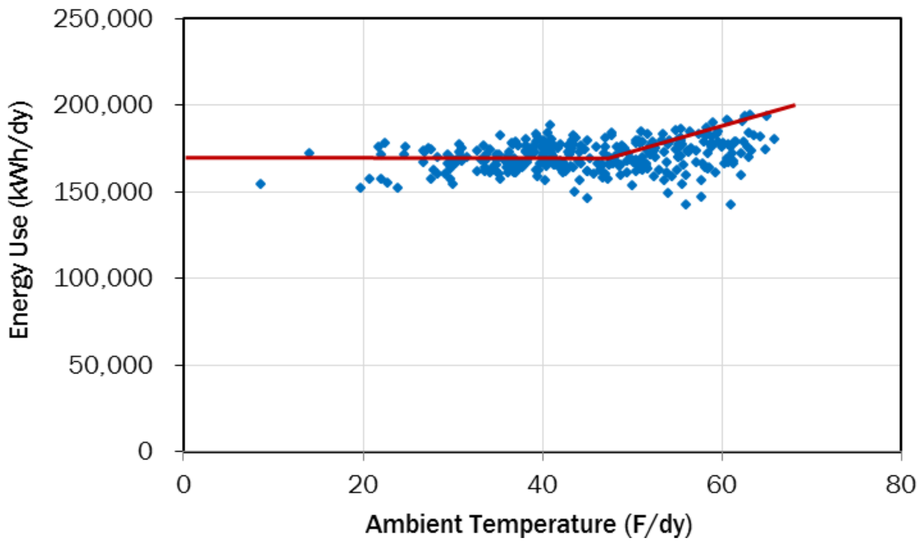


Figure 7. Example of a 3-parameter cooling change-point model.

- An R^2 that exceeds 0.7 between any two independent variables generally indicates the need to address multicollinearity.[6] A correlation matrix is useful in identifying multi-collinearity.
- Some ways to address multicollinearity include:
    - If submeters are available, split the facility into two or more measurement boundaries and split variables by measurement boundary as appropriate.
    - Re-specify the model. Consider excluding the variable that provides the least improvement to the model.
- When multicollinearity is present, the modeler should clearly explain the rationale for both the inclusion and exclusion of variables in the energy model.
- The modeler should exercise caution when excluding variables that might be significant energy drivers as this can bias the model. Further work has been done to address the effects of multicollinearity in baseline regression models by the NW Industrial Strategic Energy Management (SEM) Collaborative.[7]

2.4.1 Selection of One or Multiple Models

Some industrial facilities have distinct processes and operating modes that vary throughout the year. These may be high and low production periods such as maintenance shutdowns and seasonal production or multiple production processes that independently influence energy consumption. The resultant variation in energy use is often difficult to capture with energy drivers and indicator variables alone in a single regression model.

Developing models for the distinct operating states is a common approach to eliminate model bias between the different modes of operation. When operating modes induce a bias in the model, the use of multiple models should be considered. When the facility has one dominant mode of operation, and the energy use and expected savings during other times are small, a model that includes only this mode is generally the preferred option.

---

6 *Tools and Methods for Addressing Multicollinearity in Energy Modeling.* NW Industrial Strategic Energy Management (SEM) Collaborative. 2013.

7 *Ibid.*

Utility and end-user feedback should be solicited in the process. Judgment is required to balance accuracy versus simplicity.

Table 2. Consideration for Selection of One or Multiple Models

MODEL SELECTION	MERIT	DEMERIT
Single Model – all operational modes	• Simple to explain and use for tracking purposes. • Uses all data in the baseline period, increasing the number of observations. • Includes full range of each variable.	• Models often tend to over predict during low or no production. • R-squared values may be inflated due to extended range • Collinear variables cannot be separated to their appropriate energy meter contribution.
Single Model – one operational mode	• Model provides better prediction during production. • Eliminates the complexity of maintaining multiple models.	• Unable to estimate savings for mode(s) not modeled. • Model may not include full range of each variable.
Multiple Models	• Each model provides better prediction for all modes of operation. • Estimates savings for each mode modeled. • When applicable, separates collinear variables based on engineering Judgment of system	• Increases complexity of the tracking and measuring of energy savings. • Reduces the number of data points for each model respectively.

3. Developing a Baseline Model

3.1 Assess Statistical Significance of Independent Variables

- Screening variables for statistical significance is a critical step in the model review process, as the inclusion of erroneous variables will introduce error in the model. Likewise, the omission of critical energy driver variables will negatively affect the ability of the model to accurately characterize variation in energy use. The following guidelines can be used to test for the significance of each independent variable:
    - t-statistic > 2.0 for each variable (IPMVP 2012)[8]
    - At least one variable with a p-value < 0.10 (SEP 2017)[9]
- For ESI SEM projects, the IPMVP will serve as the official guideline.
- Appendix C shows where these values can be obtained from typical regression output tables.
- Independent variables that do not pass the above tests should not be included. Exceptions may be permissible in cases where a variable shows moderate statistical significance and is generally understood to impact energy use for the target system. The rationale for such exceptions must be documented.

8 Efficiency Evaluation Organization, Appendix B. p. 97.
9 *Superior Energy Performance Measurement and Verification Protocol for Industry.* Written under contract by The Regents of the University of California for the United States Department of Energy. March 8, 2017. Section 6.4.1 p. 23.

3.2 Statistical Criteria for Model Fitness

- While model quality cannot be judged solely on a single statistic, the fitness of the overall model can be judged against several guidelines for forecast regression models:
    - R^2: > 0.75 (IPMVP 2012)[10]
    - R^2: > 0.5 (SEP 2012)[11]
    - Net Determination Bias (NDB): < 0.005% (ASHRAE Guideline 14-2014)[12]
- For ESI SEM projects, the IPMVP will serve as the official guideline. However, the following parameters shall be documented for the overall model: R^2, adjusted R^2, coefficient of variation, NDB, auto-correlation coefficient.
- Adjusted R^2 can help determine when the addition of a variable improves the model. If adjusted R^2 decreases as variables are added, the model is likely to be over-fit.
- Appendix C shows where the basic regression parameters can be obtained from typical regression output tables.
- Plot the actual versus predicted energy use on a scatter diagram. Check that the point pattern is narrowly clustered and uniformly distributed along the diagonal as illustrated in Figure 8.

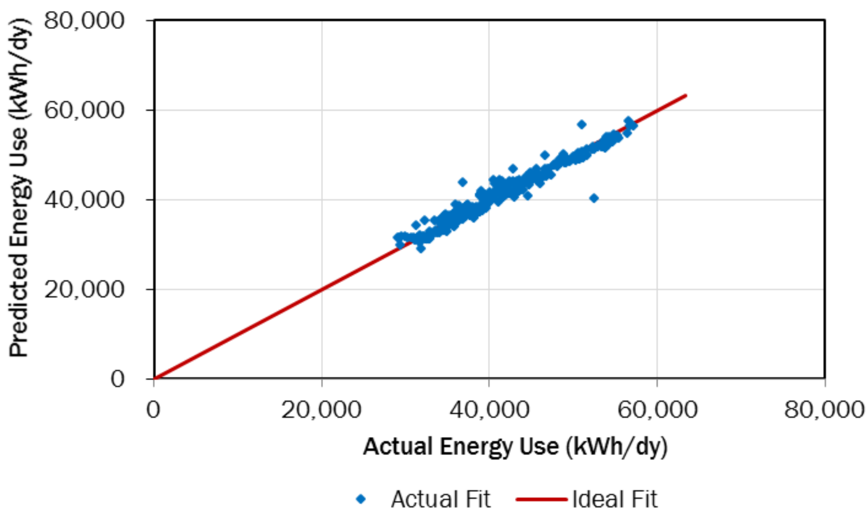


Figure 8. Example of predicted vs. actual scatter plot.

- Typically, regression-based energy models exhibit positive auto-correlation. Positive auto-correlation occurs when the sign change of the residuals is infrequent. Conversely, frequent sign changes in the residual values results in negative auto-correlation.
- There is not a defined threshold for the autocorrelation coefficient in the model development phase. However, a review of literature finds references to "light autocorrelation" for levels in the 0.3 range.[13] This becomes a factor in the uncertainty analysis, discussed in Section 4.5.1. An example of autocorrelation in a time series graph is shown in Figure 9.

---

10 Efficiency Evaluation Organization, Appendix B. p. 95.
11 The Regents of the University of California, Section 6.4.1 p. 23.
12 ASHRAE, p. 16. Table 4.2.
13 *Guidelines for Verifying Existing Building Commissioning Project Savings – Using Interval Data Energy Models: IPMVP Options B and C.* California Commissioning Collaborative. November 12, 2008. Appendix B, Page 70.

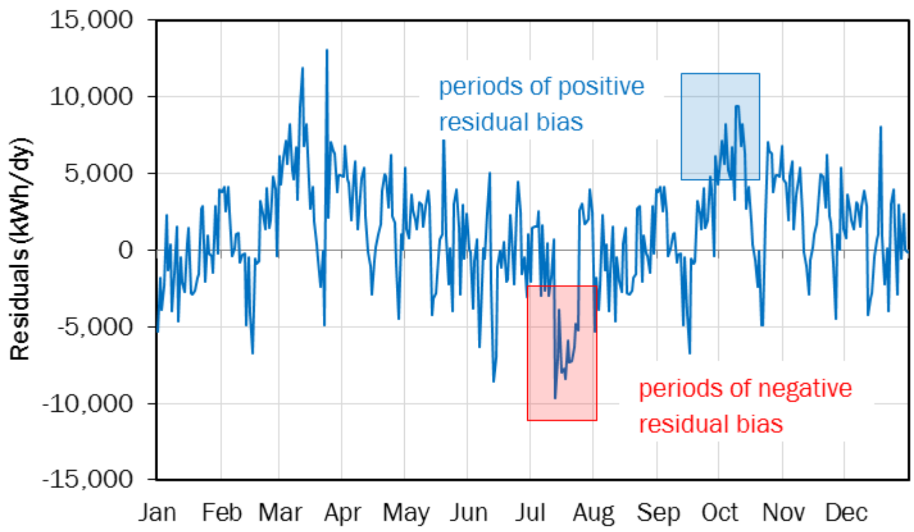


Figure 9. Example of autocorrelation in a time series graph.

- Calculate the autocorrelation coefficient (see Appendix D) and plot the model residuals over the baseline period. If autocorrelation is detected, the number of independent data points is effectively reduced. The typical remedy involves increasing the sample size, or selecting a different data interval.
- High autocorrelation may indicate the omission of a key variable, or the occurrence of an event that changed energy consumption characteristics during the baseline.
- The Durbin-Watson test can be used to determine if auto-correlation is statistically significant. The Durbin-Watson test statistic, *d*, ranges from 0-4, where:
    - $d = 2$, residuals are not correlated
    - $d \ll 2$, residuals are positively auto-correlated
    - $d \gg 2$, residuals are negatively auto-correlated
- The lower and upper bounds for the Durbin-Watson test statistic are a function of sample size, number of predictor variables, and the desired confidence level.
- The Northwest Industrial SEM Collaborative has provided a paper pertaining to autocorrelation in regression-based energy models for industrial facilities.[14]
- Residual plots that may be of value include:
    - Residuals versus time (e.g. Figure 9)
    - Residuals versus the independent variables (confirmation of homoscedastic or heteroscedastic residuals)
    - Histogram of residuals (supports Net Determination Bias)

3.3 Modifying the Hypothesis

- If the statistical tests outlined in 3.1 and 3.2 indicate insufficient fitness of the model, modify the model hypothesis. This process might include modifications to the assumed energy drivers, time intervals, change points, or the order of relationships (second order, square root, etc.).

---

14 *Tools and Methods for Addressing Autocorrelation in Energy Modeling*. NW Industrial Strategic Energy Management (SEM) Collaborative. 2013.

- If the measurement boundary is supplied by multiple meters, disaggregating the meters may result in better model resolution.
- In forming an alternative hypothesis, confirm that the characteristics of the equation remains aligned with the mechanics of the process, and that the baseline data set meets the standards outlined in Section 2.1. This information should be documented in a competing model summary. An example of a competing model summary is provided in Appendix F.

3.4 Screening for Residual Outliers

- Outliers from the residual analysis should be flagged for review. One approach for reviewing outliers is applying a common rule of thumb for identifying data that lie outside the range of ±4σ, as illustrated in Figure 10.[15] The probability that a residual will exceed ±4σ due to random chance is small. Applying a range of ±4σ eliminates unnecessary flagging of residuals, while identifying those residuals that need further review.

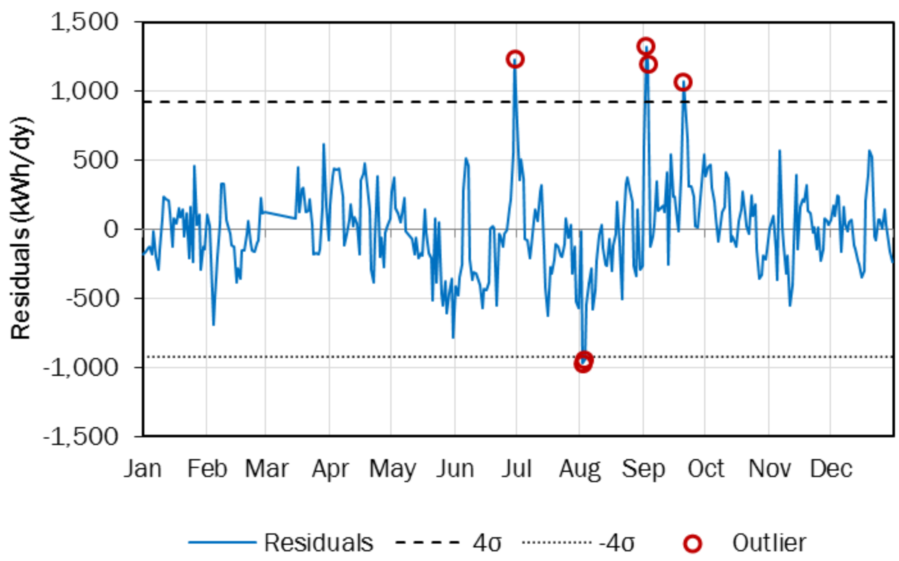


Figure 10. Inspection of residual outliers.

- Before removing outliers, the modeler should review any residuals outside the control limits with the Energy Champion to understand the cause of the anomaly.
- The modeler must provide a supporting explanation when removing statistical outliers.

3.5 Alternatives to Regression-based Forecasting

The adoption of a methodology that does not use a standard regression-based forecasting energy model may be necessary under certain conditions.

3.5.1 Backcast Approach

For the backcast approach, the regression energy model is developed from the data obtained during the reporting period. This method is applicable in instances where the resolution of the

[15] Neter, J., W. Wasserman. 1974. *Applied Linear Statistical Models.* Irwin Publishers, Homewood, Illinois. p 106.

energy data for the original baseline was relatively poor (e.g., monthly) and the resolution of the energy data during the reporting period has significantly improved.

For more details, see the Superior Energy Performance Measurement and Verification Protocol for Industry (SEP Protocol).[16]

3.5.2 Mean Model

The mean model represents the simplest form of forecasting, and may be necessary when:

- There is insufficient variation in the independent energy drivers (e.g., production is constant) such that there is also insufficient variation in the corresponding energy variable, or
- There is insufficient correlation between suspected energy drivers and energy.

For the mean model approach, the estimate of baseline energy use is the average energy use:

Baseline energy per interval = Average annual energy consumption for baseline period

This approach requires that baseline operating conditions be thoroughly documented so that changes in energy intensity observed during the reporting period can be properly assigned to EEMs directed at energy efficiency versus other changes in plant operation.

This approach is valid provided the relevant operational parameters remain within a defined range. An acceptable guideline for this tolerance is $\pm 3\sigma$ of values recorded in the baseline period. For more details, see the SEP Protocol.[17]

3.5.3 Pre-Post

For this method, a regression model is constructed using data from both the baseline and reporting period. Generally, a single indicator variable is used to estimate the difference in energy use between the two time periods, though interactive effects between energy drivers can be modeled. For more details, see the Industrial Strategic Energy Management (SEM) Impact Evaluation Report.[18]

3.6 Energy Model Report and EPT Review

The model and supporting statistics and graphics should be documented in the Energy Model Report. The EPT team will provide final approval after a review by the utility and end user.

4. Calculating Energy Savings During the Reporting Period

4.1 Maintaining Records of Events and Changes

- The savings calculated in Sections 4.3 and 4.4 represent the total (gross) energy savings for the site. In order to establish attribution, it is critical that the Energy Champion maintain accurate records of key operations and maintenance (O&M) actions or behavior-based improvements. Records of facility operations that influence energy use, including key process variables, should also be maintained. The Energy Champion should attempt to correlate inflections in the cumulative sum of differences (CUSUM) graph to these actions or changes.

16 The Regents of the University of California, Section 3.4.12, p.12.
17 The Regents of the University of California, Section 3.4.6, p.11.
18 *Industrial Strategic Energy Management (SEM) Impact Evaluation Report.* SBW Consulting, Inc. and The Cadmus Group. 2017. Appendix B., p. 61.

- Any effects from fuel switching must be accounted for and excluded from the gross MT&R savings. If fuel switching is a possibility, it is advisable to maintain records of alternate fuel sources crossing the measurement boundary beginning with the baseline period. These records can be used to document that fuel switching did not occur during the reporting period.

4.2 Adjusting for Concurrent Incentivized Projects

- If the end user is participating in other ESI program offerings, gross energy savings adjustments will likely be needed to net out savings from EEMs incentivized by other ESI components. The typical approach is an adjustment to the gross savings by the utility-approved M&V savings value associated with the project, prorated from the M&V start date to the end of the reporting period.
- Appendix B outlines the options for determining the value of the adjustment and identifying a suitable date of application.

4.3 Calculation of Savings Using Regression Model

- As data is collected during the reporting period, it should be methodically reviewed to detect anomalous values and to ensure that the independent variables fall within the ranges specified for the model. For models with a single mode of operation, the generally acceptable values for each variable will be the maximum of ±3σ or the range used in the model, as outlined in the SEP Protocol.[19] When the model includes multiple modes of operation, ±10% of the actual range is generally the most appropriate method.
- Energy savings can be calculated by applying the following equation:

$$Energy\ Savings = Predicted\ Energy\ Use - Actual\ Energy\ Use \pm Non\text{–}\ Routine\ Adjustments$$

- For periods with infrequent occurrences of out-of-range variables, the magnitude of energy savings should be reviewed. Generally, no further adjustments are needed if energy savings are similar to the other observations within the ranges specified by the model.
- When variables exceed the valid range of the model, capping production variables may be necessary to avoid overestimating energy savings. If capping is applied, all values must be capped consistently.
- If an acceptable capping limit cannot be determined, an expected value of energy savings may be provided. If an expected value cannot be determined, then energy savings for these occurrences should be excluded.
- Occurrences of abnormal energy savings, i.e., exceeding ± 4σ, should be reviewed. Plant operations can be reviewed with the Energy Champion if further questions persist upon reviewing the data. The expected or average value of savings can be used for these anomalous observations.
- The CUSUM calculation is an effective means of quantifying the total energy savings benefit. In graphical form, the CUSUM provides a powerful illustration of the total savings achieved during a specified reporting period. However, the CUSUM graph should be used in conjunction with a time series plot of energy and the independent variables. Together, these graphs help establish an informed understanding of energy intensity inflections. An example of a CUSUM graph is shown in Figure 11.

[19] The Regents of the University of California, Section 3.4.6, p.12.

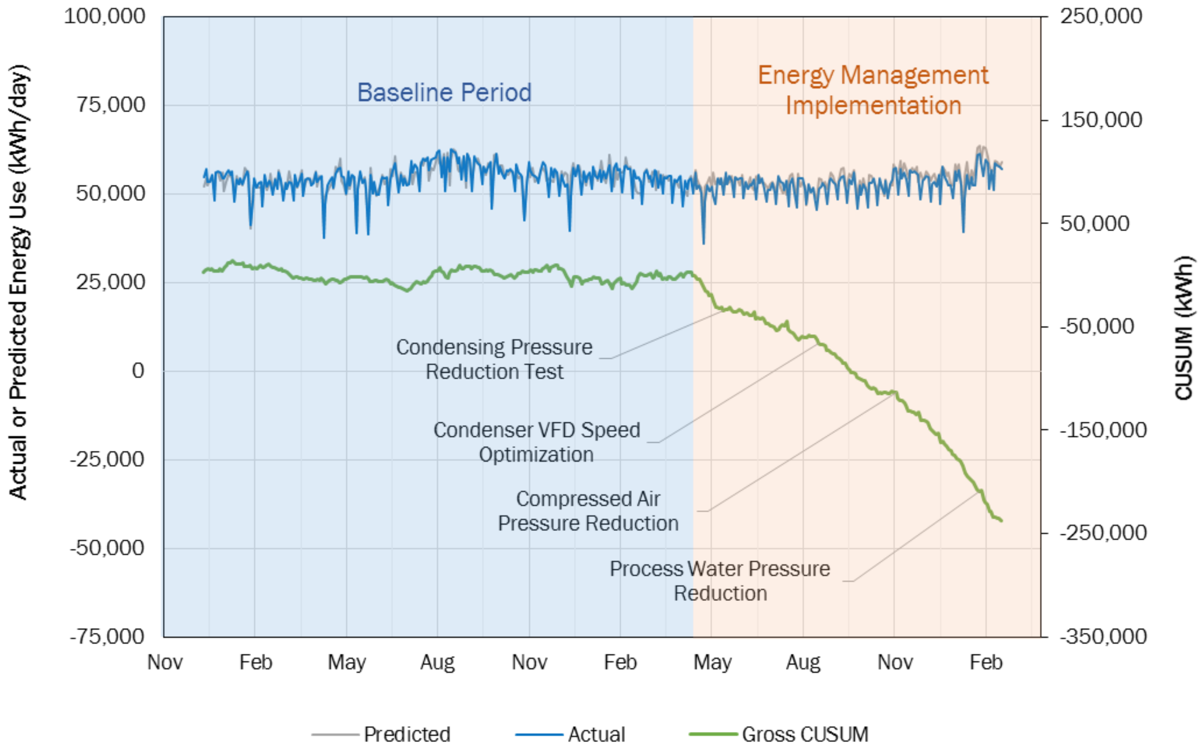


Figure 11. CUSUM graph example.

4.4 Calculation of Savings Using Alternative Approaches

4.4.1 Savings Calculation by Backcast Approach

When using the backcast approach, separate energy models are created for each reporting period. Each respective model estimates energy use during the baseline period using the weather and production observed during the baseline period. A timeline for the back-casting procedure is illustrated in Figure 12.

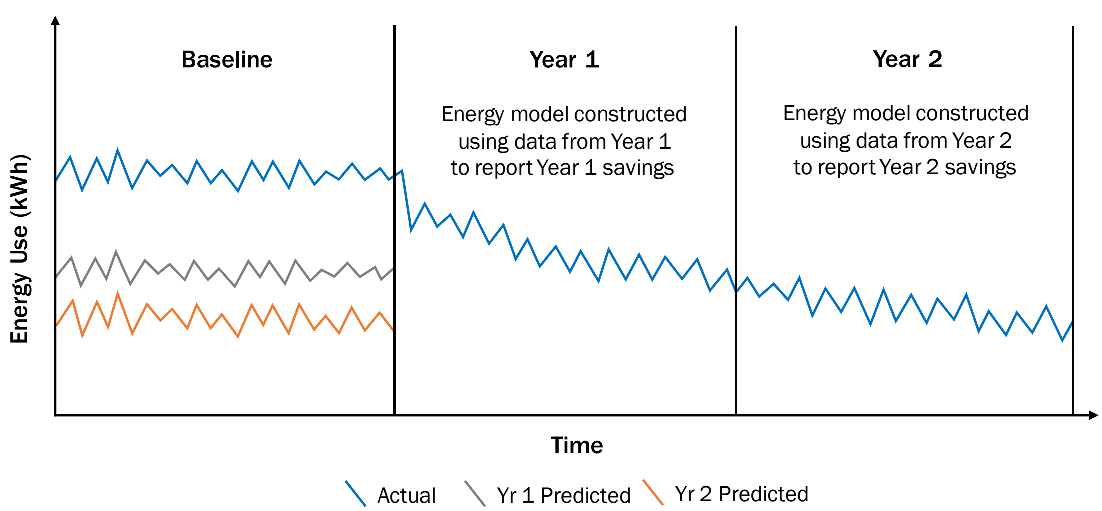


Figure 12. Backcast approach.

To calculate energy savings for Year 1, first an energy model is created using actual energy, weather, and production data from Year 1. This model is then used to predict energy use during the baseline period based on weather and production data reported during that same baseline period. Finally, savings are calculated using the actual energy use during the baseline period and the energy use predicted for the baseline period using the Year 1 model.

Thus, energy savings for the Year 1 reporting period are calculated as:

$$\begin{aligned} Energy\ Savings_{Year\ 1} \\ &= (Actual\ Energy\ Use)_{Baseline} \\ &- \ (Predicted\ Energy\ Use, Year\ 1\ Model)_{Baseline} \\ &\pm\ Non\text{–}\ Routine\ Adjustments \end{aligned}$$

Likewise, the energy savings for the Year 2 reporting period are based on the model created using energy use, weather, and production data from Year 2 and the energy use, weather, and production reported during the baseline. Energy savings for the Year 2 reporting period are calculated as:

$$\begin{aligned} Energy\ Savings_{Year\ 2} \\ &= (Actual\ Energy\ Use)_{Baseline} \\ &- \ (Predicted\ Energy\ Use, Year\ 2\ Model)_{Baseline} \\ &\pm\ Non\text{–}\ Routine\ Adjustments \end{aligned}$$

4.4.2 Savings Calculation by Mean Model

For a mean model, baseline energy is calculated as the mean or average energy use during the baseline period. For a given time interval, energy savings are then calculated as the difference between the mean value from the baseline period and the actual energy use for that time interval, plus or minus any adjustments.

$$\begin{aligned} Energy\ Savings = Mean\ (Actual\ Energy\ Use)_{Baseline} - (Actual\ Energy\ Use)_{Reporting} \\ \pm\ Non\text{–}\ Routine\ Adjustments \end{aligned}$$

4.4.3 Savings Calculation by Pre-Post Approach

For models with a single indicator variable, the savings estimate per time interval is the estimated coefficient of the indicator variable. The Industrial Strategic Energy Management (SEM) Impact Evaluation Report provides more details for calculating energy savings when the indicator variable (for the reporting period) is included as an interaction term with other model variables.[20]

4.4.4 Savings Calculation by Bottom-up Approach

Quantification of energy savings using a bottom-up approach consists of engineering calculations supported by short-term data logging. The application of this approach is limited to specific cases when top-down, whole-facility energy modeling efforts are unsuccessful. This approach may also be used for comparison purposes. Further information regarding the application of engineering calculations including determination of the baseline, calculations of energy savings, and required

20 SBW Consulting, Inc. and The Cadmus Group, Appendix B, p. 73.

project documentation is provided in BPA's Engineering Calculations with Verification (ECwV) Protocol.[21]

4.5 Options for Establishing Statistical Confidence of Savings Value

4.5.1 Uncertainty in the Forecasting Estimate

In certain instances, it may be necessary to specify a range of energy savings performance for a defined statistical confidence level.

ASHRAE provides a detailed description of uncertainty analysis.[22] The following methodology provides an approach for calculating uncertainty derived from model error. This method is a simplified version of the uncertainty analysis provided in the Industrial Strategic Energy Management (SEM) Impact Evaluation Report.[23] It should be noted that this approach does not capture error associated with measurement hardware. In most cases, the measurement error component should be small relative to the regression model error.

The fractional savings uncertainty (FSU) for the majority of ESI MT&R models can be estimated by the following equation:

$$FSU = 1.26t \times \frac{CV\left[\left(\frac{n}{n'}\right)\left(1+\frac{2}{n'}\right)\left(\frac{1}{m}\right)\right]^{\frac{1}{2}}}{F}$$

Where:

t	=	*t*-statistic for desired confidence level
CV	=	coefficient of variation
n	=	number of observations in the baseline period
m	=	number of observations in the reporting period
F	=	fractional savings

The effective number of observations in the baseline period, n', after accounting for auto correlation is:

$$n' = n\frac{(1-\rho)}{(1+\rho)}$$

Where:

ρ	=	auto-correlation coefficient

While the preceding methodology is generally applied to analyze savings uncertainty in an ex-post analysis, this analysis can be used to inform the model development, particularly when the model developer is faced with multiple options related to time interval or variable selection.

21 *Engineering Calculations with Verification Protocol, Version 1.0.* Bonneville Power Administration. 2012. www.bpa.gov/EE/Policy/IManual/Documents/July%20documents/6_BPA_MV_ECwV_Protocol_May2012_FINAL.pdf.
22 ASHRAE, Annex B.
23 SBW Consulting, Inc. and The Cadmus Group, Appendix B, p. 75.

4.5.2 Statistical Confidence for Backcast Method

The FSU equation can also be used to estimate savings uncertainty for the backcast method. When using the FSU equation, the model statistics and "baseline" observations (n) occur during the reporting period of the project. Likewise, the number of observations during the "reporting" period (m) occur during the baseline period of the project.

4.5.3 Statistical Confidence for Mean Model

When applying the mean model approach, two-sided t-tests are performed on energy use and assumed energy drivers prior to reporting energy savings. The t-test should demonstrate that the energy use of the reporting period is less than the baseline period. It must be shown that assumed energy drivers did not influence energy savings. T-tests or other methods may be used to demonstrate this. All t-tests should be performed at the 80% level of confidence using methods for equal or unequal variances as appropriate for the samples under study.

4.5.4 Statistical Confidence for Pre-Post

When using the pre-post method, the indicator variable's standard error is used to determine the uncertainty of the savings estimate. For a desired level of confidence, the t-stat or p-value can be used to determine the confidence in the savings estimates.

4.6 EPT Review and Approval

The savings calculation methodology and verified savings value will be documented in the SEM Completion Report. The EPT team will provide final sign-off, but BPA's Energy Management Engineering COTR (E-COTR) will provide final authorization of the savings and incentive.

5. Making Non-Routine Adjustments

5.1 Scenarios for Model Reassessment

- The model is considered valid for the range of the independent variables observed during the baseline period, provided the general operation and qualitative factors of the facility or system remain constant throughout the reporting periods. The SEP Protocol provides an additional provision that validates the model if the independent variable is within $\pm 3\sigma$ from the mean of the baseline data set.[24]
- Non-programmatic effects may occur during the reporting period. Such scenarios would trigger a reassessment of the energy model. These scenarios can be characterized into three different categories of increasing complexity:

5.1.1 Static Change Assessment

A static change is a change in electric load within a well-defined boundary and with minimal interactive effects. Examples of a static change are:

- Addition of a new exhaust fan for safety/environmental purposes
- Added section of the facility in which the energy flows can be easily isolated

[24] The Regents of the University of California, Section 3.4.6, p.11.

5.1.2 Minor Process Change Assessment

A minor process change is a distinct change in operations that does not fundamentally change the process itself. These changes generally impact one or just a few production or process variables. Examples of a minor process change are:

- Change in business operations that requires a new independent variable (e.g., new product type)
- Change in the operating pressure of a sub-system within the plant

5.1.3 Major Process Change Assessment

A major process change affects the fundamental energy consumption characteristics of the facility, rendering the original model specification invalid. These changes may impact many systems within the plant. Examples of a major process change are:

- A sustained increase or decrease in the observed level of an independent variable outside the range for which the baseline model was established.
- A change in plant operations from batch-type to continuous

5.2 Options for Baseline Adjustment

Baseline adjustments should reflect the scenario encountered.

5.2.1 Static Change Adjustment

The change in electrical load should be accounted for based on sub-metered data and accompanying analysis.

- For constant loads, annual energy use can often be extrapolated using short-term (e.g. two weeks') data logging.
- For variable loads, long-term or permanent submetering is preferred. Where long-term submetering is not feasible, empirical models can be developed that correlate energy use from these loads to weather, production and/or process variables.
- For relatively small static changes, engineering calculations supported with motor nameplate information may be acceptable.

5.2.2 Minor Process Change Adjustment

To account for a minor process change, a regression approach is generally preferred. The model must include sufficient data before and after the change to accurately estimate the impact of this change. Production or process data is required for documentation of when this change occurred.

- When the change is an added product, a regression model, including the added product, can be used to estimate the change in energy use for this product. Generally, the other variables are the same variables used in the energy model. The estimated coefficient of the new variable can then be added to the energy model.
- When a change in sub-system operation occurs, a regression model with an indicator variable can be evaluated. Again, the other variables are the same variables used in the energy model and the indicator variable is set to one when the change occurs. The estimated coefficient of the indicator variable can then be added to the energy model.
- When the regression model is not a suitable approach, estimates of the change may be made based on engineering calculations or published data.

5.2.3 Major Process Change Adjustment

Like minor process changes, a regression approach is preferred.

- When the process itself has fundamentally changed, creating a new regression model or re-baselining may be necessary. Consideration of the implementation dates of the EEMs need to be considered when changing the time period of the model.
- When independent variables are frequently outside the acceptable limits of the model, a new regression model may be required. The SEP Protocol provides a “chaining adjustment” methodology to model these situations.[25]
- Other options for dealing with a major process change include a pre-post or bottom-up approach.

5.3 Guidelines for Modification of Regression Model

When revising the baseline model is necessary, the revised baseline period must adequately capture the new range of operating conditions, including seasonal cycles (if applicable). Until a new model can be established, SEM savings incentives would typically be put on hold but the accumulated savings that preceded the retrofit would be considered based on engineering calculations with verification.

5.4 EPT Approval

When a baseline model must be adjusted, the proposed adjustment should be reviewed and approved by the EPT team in advance of any modeling work.

6. Projecting Energy Savings from a Condensed Performance Period

The following section outlines four methods to project annual energy savings if less than a year of data is available during the performance period. Under the current SEM program, this method would seldom be applicable. However, in the case of meter failure or other unforeseen circumstances, these methods, which were developed and tested for Track and Tune projects commencing prior to October 1, 2015, may be applicable. The projected Year 1 energy savings are based on the achieved energy savings obtained during the performance period, which is typically 90 days.

For each of these methods, it is essential that the following factors are taken into account:

- The number of valid observations during the performance period.
- The expected number of valid observations during the remainder of Year 1.
- The expected distribution of the energy drivers during the remainder of Year 1 relative to the distribution of the energy drivers during the performance period.
- Engineering and program Judgment on the likelihood of savings to persist.

6.1 Direct Percentage Basis

When the distribution of the energy drivers is expected to be the same for the remainder of Year 1, Year 1 energy savings can be projected by extrapolating percent energy savings from the performance period.

[25] The Regents of the University of California, Section 3.6.5, p.17.

6.2 Percentage Basis with Forecast of Energy Drivers

When the distribution of energy drivers is expected to be different for the remainder of Year 1, the distribution of energy drivers must be considered when projecting Year 1 energy savings. For example, if during the performance period, energy savings were only obtained when production was low, then the expected distribution of production should be used to project Year 1 energy savings. If production is expected to be high for the majority of Year 1, it would be incorrect to project Year 1 savings based on savings achieved during the performance period when production was low.

6.3 Normalized Annual Consumption

- This method can be used in lieu of the "Percentage Basis with Forecast of Energy Drivers" method described above (Section 6.2). This method requires the development of a second regression model for the performance period. The total derivative of the baseline energy equation is taken to develop a governing equation. The inputs for the governing equation are the coefficients from the baseline and performance period models, as well as the projected distribution of energy drivers. TMY3 weather data is typically used for weather dependent energy drivers and the best estimate of Year 1 production is used for production energy drivers.
- This modeling approach provides a disaggregation of energy savings by energy drivers, which provides transparency for how energy savings were achieved.
- The weaknesses of this approach are that it requires additional calculation steps and that the energy signature of the baseline and performance periods must be the same.
- This method is similar to the Standard Condition Adjustment Model defined by SEP.[26]

6.4 Pre-Post

- This method can be used in lieu of the "Direct Percentage Basis" method described in Section 6.1. This method was used by Cadmus for the 2012 and 2017 Energy Management Impact Evaluation, and follows a methodology described by Luneski (2011).[27] This method entails developing a new regression model using an indicator variable to differentiate the baseline and performance period data. The value of the indicator variable represents the energy savings.
- This modeling approach does not normalize the savings value for annual weather or production and thus it should not be used when the distribution of the energy drivers is expected to be significantly different for the remainder of Year 1.

[26] The Regents of the University of California, Section 6.2.3, p.19.

[27] Luneski, R.D. 2011. *A Generalized Method for Estimation of Industrial Energy Savings from Capital and Behavior Programs.* Industrial Energy Analysis.

Appendix A – Treatment of EEMs During the Baseline Period

DESCRIPTION (IN ORDER OF PREFERENCE)	GUIDELINES	MERITS	DEMERITS
Standard Approach Select a baseline period without capital projects and immediately prior to the reporting period. $y\left(\frac{kWh}{period}\right) = \beta_0 + \beta_1 x_1 + \beta_i x_i$	• Verify absence of utility-incentivized EEMs by interviewing facility and speaking to serving utility. • Confirm energy intensity profile is consistent over the selected period.	• Incorporates the full data set in the baseline model. • Requires no manipulation of data. • Requires no adjustments during reporting period.	• No obvious demerits, provided energy intensity profile is consistent throughout baseline period.
Year-End MT&R Adjustment Choose a baseline period immediately prior to the first capital project. Subtract M&V savings from <u>year-end</u> MT&R savings. $y\left(\frac{kWh}{period}\right) = \beta_0 + \beta_1 x_1 + \beta_i x_i + (IV = 0,1)_K (M\&V)_K$	• Maximum exclusion period = 12 months. • Exclusion period must have a consistent energy profile, aside from the EEM(s).	• Provides direct reconciliation with EEM M&V value. • Requires no adjustment of baseline data set.	• Data immediately preceding reporting period is excluded. • M&V adjustment must be performed throughout reporting period.
Pre-EEM Baseline Normalization by M&V Value Adjust the pre-EEM baseline values by the EEM M&V value. $y\left(\frac{kWh}{period}\right) = \beta_0 + \beta_1 x_1 + \beta_i x_i$	• EEM completion report must be reviewed and included as attachment. • Interactive effects described in project report must be factored in to baseline adjustment.	• Provides direct reconciliation with M&V value. • Enables use of the entire baseline data set. • CUSUM for reporting period starts at zero.	• Requires adjustment to baseline data set (IPMVP does not prohibit). • Accurately incorporating interactive effects is challenging and labor intensive.
Baseline Normalization by Factored Indicator Variable Apply an indicator variable in the baseline data set, representing the implementation of an EEM. The indicator variable may or may not be factored with one or more primary independent variables to account for interactive effects. $y\left(\frac{kWh}{period}\right) = \beta_0 + \beta_1 x_1 + \beta_i x_i + \beta'(IV = 0,1)x'$	• Factored indicator variable will add to the number of points required in the baseline data set (*n* × 6).	• Allows regression model to solve for interactive effects of EEM with other energy drivers. • Yields the highest R^2.	• No reconciliation with EEM's M&V value. • If backsliding occurred on the EEM, program component would pick up any recapturing of the original savings.
Indicator Variable Representation of Non-Incentivized EEM To prevent incentivizing a previously implemented non-incentivized EEM by program component, apply an indicator variable representing implementation of the EEM then solve for the coefficient. $y\left(\frac{kWh}{period}\right) = \beta_0 + \beta_1 x_1 + \beta_i x_i + \beta'(IV = 0,1)x'$ **Describes an independent scenario*	• Non-incentivized EEMs implemented during baseline period should be accurately reflected in baseline model.	• Prevents "free-rider" EEMs from inflating the savings associated with program component. • Allows use of the entire baseline data set.	• The quantification of the savings associated with the EEM is limited to the precision of the model.

Appendix B – Treatment of Incentivized EEMs Installed During the Reporting Period

<table>
<tr><th rowspan="2">PROJECT INSTALLED</th><th rowspan="2">SAVINGS OBSERVED IN CUSUM?</th><th rowspan="2">M&V STATUS</th><th colspan="2">PRORATING METHOD</th></tr>
<tr><th>START DATE</th><th>SAVINGS VALUE</th></tr>
<tr><td>No or Incomplete</td><td>n/a</td><td>n/a</td><td>n/a</td><td>n/a</td></tr>
<tr><td rowspan="10">Yes</td><td rowspan="3">No</td><td>Not started</td><td>n/a</td><td>n/a</td></tr>
<tr><td>In progress</td><td>Use the actual project M&V start date.</td><td>Wait for M&V to be completed (if an early estimate is needed, solve for value in CUSUM).</td></tr>
<tr><td>Completed</td><td>Use the actual project M&V start date.</td><td>Use site savings M&V value.</td></tr>
<tr><td rowspan="7">Yes</td><td rowspan="3">Not started</td><td rowspan="3">Based on CUSUM inflection and ideally supported by email from ESIP (e.g., equipment was commissioned on xx/xx date).</td><td>Option A. Solve for savings value using indicator variable during reporting period.</td></tr>
<tr><td>Option B. Use estimated site savings from custom project proposal.</td></tr>
<tr><td>Option C. If the savings value from A and B differ significantly, confer with EPT team.</td></tr>
<tr><td rowspan="2">In progress</td><td>Option A. Based on CUSUM inflection, and ideally supported by email from ESIP.</td><td rowspan="2">Wait for M&V to complete (if an early estimate is needed, solve for value).</td></tr>
<tr><td>Option B. At the latest, use Actual Project M&V Start Date.</td></tr>
<tr><td rowspan="2">Completed</td><td>Option A. Based on CUSUM inflection and ideally supported by email from ESIP.</td><td rowspan="2">Use site savings M&V value.</td></tr>
<tr><td>Option B. At the latest, use Actual Project M&V Start Date.</td></tr>
</table>

Appendix C – Overview of Regression Output

```
Baseline relationship for Production Days Only

m(formula = Total_KWH ~ IND_early + IND_late + IND_missingkWh +
    Prod_carrots + Prod_Corn + Prod_Peas + WetBulb_KHRI, data =
Dataset)

Residuals:
   Min      1Q Median     3Q    Max
-38223  -7100    358   8095  32761

Coefficients:
                 Estimate Std. Error t value Pr(>|t|)
(Intercept)     2.038e+04  9.919e+03   2.054   0.0416 *
IND_early      -5.203e+04  3.998e+03 -13.012  < 2e-16 ***
IND_late        4.889e+04  3.998e+03  12.229  < 2e-16 ***
IND_missingkWh -2.515e+04  6.204e+03  -4.054 7.97e-05 ***
Prod_carrots    9.017e-02  7.928e-03  11.373  < 2e-16 ***
Prod_Corn       8.252e-02  5.217e-03  15.819  < 2e-16 ***
Prod_Peas       6.696e-02  5.122e-03  13.075  < 2e-16 ***
WetBulb_KHRI    6.573e+02  1.596e+02   4.120 6.18e-05 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 13170 on 154 degrees of freedom
Multiple R-squared: 0.8452, Adjusted R-squared: 0.8381
F-statistic: 120.1 on 7 and 154 DF,  p-value: < 2.2e-16
```

Figure C-1. Regression output from "R" open source statistical software.

Regression Statistics	
Multiple R	0.965375
R Square	0.931949
Adjusted R Square	0.916827
Standard Error	590.4573
Observations	12

CV =	0.03

Note: CV must be calculated separately.

ANOVA

	df	*SS*	*MS*	*F*	*ignificance F*
Regression	2	42971374	21485687	61.627181	5.59E-06
Residual	9	3137758	348639.8		
Total	11	46109132			

	Coefficients	*andard Err*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*ower 95.0%*	*Ipper 95.0%*
Intercept	24373.82	5738.171	4.247664	0.0021499	11393.18	37354.46	11393.18	37354.46
Ave Temp	-428.045	208.7444	-2.05057	0.0705487	-900.258	44.1677	-900.258	44.1677
Ave Temp^2	5.392718	1.845353	2.922323	0.0169676	1.218239	9.567196	1.218239	9.567196

Figure C-2. Regression output from Microsoft Excel.

Appendix D – Glossary of Terms

The definitions below address terms used within the body of this document, presented in the context of ESI's MT&R procedure. For a more comprehensive overview of statistical terms related to measurement and verification, please refer to BPA's Glossary for M&V: Reference Guide.[28]

1. Adjusted R^2: A measure of the total variation accounted for in the model that penalizes for the number of parameters used in the model.

2. Autocorrelation Coefficient: A measure of the correlation of a time series with its past and future values (also referred to as serial correlation). In a time series plot of residuals, autocorrelation is characterized by a tendency for the bias in data point n to be a predictor of a similar bias in data point $n + 1$. The autocorrelation coefficient can be calculated by performing regression on two identical data sets, offset by one unit of time. The square root of the resulting coefficient of determination is the autocorrelation coefficient (ρ) for the data set.

   Auto-correlation can also be calculated from the residuals, e, from the following equation:

$$\rho = \frac{\sum_{t=2}^{n} e_t \ e_{t-1}}{\sum_{t=1}^{n} e_t^2}$$

3. Change-Point Model: A model in which the relationship of a dependent variable is discontinuous with respect to an independent variable. The change-point is the value of the independent variable at which this discontinuity occurs. In the context of industrial energy efficiency, a common scenario arises when the energy intensity of a building or system changes at a specific ambient temperature, at which the HVAC system switches from heating mode to cooling mode.

4. Coefficient of Determination (R^2): Statistically, the proportion of the total variation in the dependent variable that is explained by the regression equation. Mathematically, defined as

$$R^2 = \frac{\sum(\hat{Y}_i - \overline{Y})^2}{\sum(Y_i - \overline{Y})^2},$$

   where,
   - $\hat{Y}_i$ = the predicted energy value for a particular data point using the measured value of the independent variable.
   - $\overline{Y}$ = mean of the n measured energy values, $\overline{Y} = \frac{\sum Y_i}{n}$.
   - Y_i = actual observed value of the dependent variable.

5. Coefficient of Variation (CV RMSE): The CV is calculated as the ratio of the root mean squared error (RMSE) to the mean of the dependent variable (energy). CV is a dimensionless value, and the ratio is typically multiplied by 100 and given as a percentage. CV aims to describe the model fit in terms of the relative sizes of the squared residuals. CV evaluates the relative closeness of the predictions of the actual values (the uncertainty of the model), while R^2 evaluates how much of the variability in the actual values is explained by the model.

[28] *Bonneville Power Administration's Glossary for M&V: Reference Guide, Version 1.1.* Bonneville Power Administration. May 2012.

$$CV\ (RMSE) = \frac{\sqrt{\left(\frac{\sum(\hat{y}_i - y_i)^2}{(n-p)}\right)}}{\bar{y}} \times 100$$

6. Cooling Degree Days (CDD): A measure of how many degrees the outside air temperature (T_{oa}) is above the balance point (T_{bal}) over the course of a day. The units CDD are °F · days. When using average values of T_{oa}, CDD can be calculated as:

$$\mathrm{CDD}(T_{bal}) = 1\mathrm{day} \times \sum_{n=1}^{days} (T_{oa}, n - T_{bal})+$$ [29]

Note that different time intervals can lead to different values for degree-days. A source for degree days is: www.degreedays.net.

7. Data Champion: This person, assigned by the end user, is the point of contact for data review and collection. This person may be the Energy Champion or report to the Energy Champion.

8. Energy Champion: This person, assigned by the end user, determines potential energy efficiency projects and tracking techniques.

9. Energy Efficiency Measure (EEM): Equipment and/or actions taken to reduce electrical energy use.

10. Fractional Savings Uncertainty (FSU): The calculated uncertainty in the total savings over *m* time periods divided by the total savings over the same time period, where uncertainty is measured as the quantity of savings from the upper confidence limit to the lower confidence limit surrounding a savings estimate.

11. Heteroscedasticity: In contrast to homoscedasticity, this occurs when error (or residual) variance is not constant throughout the observations e.g., when the residual variance is shown to increase or decrease with the value of an independent variable.

12. Heating Degree Days (HDD): A measure of how many degrees the outside air temperature (T_{oa}) is below the balance point (T_{bal}) over the course of a day. The units HDD are °F · days. When using average values of T_{oa}, HDD can be calculated as:

$$\mathrm{HDD}(T_{bal}) = 1\mathrm{day} \times \sum_{n=1}^{days} (T_{bal} - T_{oa}, n)+$$ [30]

Note that different time intervals can lead to different values for degree-days. A source for degree days is: www.degreedays.net.

13. Homoscedasticity: Homoscedasticity generally means that all data in a model have similar variance over the modeling period. Within linear regression, this means that the variance around the regression line is similar for all values of the dependent variables.

29 Kreider, Curtiss, Rabl. 2002. *Heating and Cooling of Buildings, Second Edition.* McGraw Hill. p. 381.
30 ibid, p. 379.

14. Indicator Variable: (Also referred to a categorical variable.) A variable used to account for discrete levels of a qualitative variable. Generally, indicator variables are assigned a value of 0 or 1 to account for different modes of operations, and a qualitative variable with *r* levels can be modeled with *r* – 1 indicator variables.

15. International Measurement and Verification Protocol (IPMVP): The IPMVP provides an overview of current best practice techniques for verifying results of energy efficiency, water efficiency, and renewable energy projects in commercial and industrial facilities. It may also be used by facility operators to assess and improve facility performance. The IPMVP is the leading international standard in Measurement and Verification protocols.[31]

16. Measurement and Verification (M&V): The process of planning, measuring, collecting and analyzing data for the purpose of verifying and reporting savings within an individual facility resulting from the implementation of EEMs.[32]

17. Measurement Boundary: A notional boundary drawn around equipment and/or systems to segregate those which are relevant to savings determination from those which are not. All energy uses of equipment or systems within the measurement boundary must be measured or estimated, whether the energy uses are within the boundary or not.

18. Mean Model: (Also referred to as a single parameter model.) A model that estimates the mean of the dependent variable.

19. Monitoring, Tracking, and Reporting (MT&R): MT&R refers to the measurement systems, statistical tools, and business practices associated with measuring energy intensity, establishing targets for improvement, and reporting results and impacts. MT&R has many similarities to the Plan-Do-Check-Act (PDCA) methodology that is central to several widely adopted business performance standards.

20. Multicollinearity: A phenomenon in which two or more independent variables in a multiple regression model are correlated.

21. Net Determination Bias Error (NDB or NBE): A statistical metric that quantifies the tendency of a model to underestimate or overestimate savings. Typically represented as a percentage. Note that if regression is performed properly, net determination bias should be zero. A positive value indicates a tendency of the model to overestimate savings. NDB is calculated as:

$$NDB = \frac{\sum(Y_i - \hat{Y}_i)}{\sum Y_i} \times 100$$

22. Non-programmatic Effects: Factors that did not occur during the baseline period and are outside the influence of the program.

23. Regression Model: A mathematical model based on statistical analysis where the dependent variable is regressed on the independent variables which are said to determine its value. In so doing, the relationship between the variables is estimated statistically from the source data.

---

[31] Efficiency Evaluation Organization.

[32] *Ibid.*

24. Strategic Energy Management (SEM): The application of the business principles of continuous improvement to drive systematic, long-term reductions in the energy intensity of a system, facility, or organization.

25. Tune-up: A major on-site technical effort, led by a tune-up engineer, which may result in immediate operational changes and a prioritized list of low-cost/no-cost action items.

Appendix E – Models with Irregular Time Intervals

When developing an energy model based on data of varying intervals, time intervals must be accounted for in the regression analysis or the model will be biased. This is accomplished by first converting the data for each observation of the independent and response variables to average values. Then all dependent and independent variables need to be weighted by the number of intervals in the billing period. This can be accomplished by using weighted regression analysis, or duplicating each observation by the number of time intervals in the billing period.

Energy models with irregular time intervals occur most often when developing energy models with monthly utility bills. Consider, for example, the case when the billing period for each utility bill is different. When developing the energy model, the model must account for this irregular time interval to eliminate bias from the varying time periods. Table E-1. shows the data per billing period and the daily average values for this data. Note that because T_{db} was already provided as an average value, this value is the same for both the billing period and the daily average.

Table E-1. Example data set for weighted regression

Billing Period					Daily Average		
Billing Period	**Days/Billing Period**	**Electricity Use (kWh/Billing Period)**	**Avg. Tdb (°F/Billing Period)**	**Production (lbs/Billing Period)**	**Electricity Use (kWh/dy)**	**Avg. Tdb (°F/dy)**	**Avg. Production (lbs/dy)**
Jan	27	227,772	39.0	2,649	8,436	39.0	98.1
Feb	29	246,471	39.7	2,448	8,499	39.7	84.4
Mar	28	142,072	42.1	2,335	5,074	42.1	83.4
Apr	29	172,318	48.2	1,891	5,942	48.2	65.2
May	28	123,368	52.5	1,229	4,406	52.5	43.9
Jun	39	126,945	61.3	1,685	3,255	61.3	43.2
Jul	29	101,529	66.8	1,595	3,501	66.8	55.0
Aug	29	133,429	67.4	2,042	4,601	67.4	70.4
Sep	33	150,975	63.5	2,290	4,575	63.5	69.4
Oct	30	144,720	52.7	2,112	4,824	52.7	70.4
Nov	24	140,880	47.5	1,596	5,870	47.5	66.5
Dec	38	221,502	37.4	1,661	5,829	37.4	43.7
Total/Avg.	**363**	**1,931,981**	**51.5**	**1,961**	**5,401**	**51.5**	**66.1**

After the average values per interval are obtained, in this case daily average values, the analysis can be performed by using weighted regression or duplicating each observation by the corresponding number of time intervals for each observation. When using weighted regression, the weights, W, correspond to the number of time intervals per observation. For this example, the diagonal matrix W_{ii} would be:

$$W_{ii} = [27, 29, 28, 29, 28, 39, 29, 29, 33, 30, 24, 38]$$

When duplicating observations, each observation of average values is duplicated by the number of time intervals for the observation. In this example, the observations for January would be duplicated 27 times; the observations for February would be duplicated 29 times, and so forth. A spreadsheet can be used to facilitate duplicating observations.

A weighted regression set is developed to demonstrate how weighted regression is performed by duplicating observations as described above. Then both the weighted regression set and the daily average, or ordinary least squares regression set, is fit to a three-parameter, multivariable heating model as:

$$Energy\ Use\ \left(\frac{kWh}{dy}\right) = \beta_o + \beta_1(\beta_2 - Avg.Daily\ Temp)^+ + \beta_2(Avg.Daily\ Saw\ Dust)$$

Table E-2 shows that the regression coefficients calculated using weighted regression are different from the ordinary least squares method.

Table E-2. Coefficient results from weighted and ordinary regression analysis

	Weighted (Observations = 363)	Ordinary (Observations = 12)
Bo	1,477.6960	1,518.1765
B1	124.4626	125.1822
B2	58.5320	58.5860
B3	42.1438	41.4257

Table E-3 shows that the sum of the residuals for ordinary regression analysis differs from zero. This difference is caused by bias in the model coefficients. The sum of the residuals for weighted regression is nearly zero. This difference of -1 is the result of numerical errors in transferring coefficient values from the modeling program to the calculation spreadsheet and underscores the necessity of reporting and using coefficients with adequate precision.

Table E-3. Comparison of residuals between weighted and ordinary regression analysis

Actual		Weighted		Ordinary	
Billing Period	**Electricity Use (kWh/Billing Period)**	**Predicted Electricity Use (kWh/Billing Period)**	**Residual (kWh/Billing Period)**	**Predicted Electricity Use (kWh/Billing Period)**	**Residual (kWh/Billing Period)**
Jan	227,772	217,161	10,611	216,914	10,858
Feb	246,471	213,977	32,494	213,982	32,489
Mar	142,072	197,054	-54,982	197,031	-54,959
Apr	172,318	159,831	12,487	160,059	12,259
May	123,368	114,200	9,168	114,761	8,607
Jun	126,945	128,634	-1,689	129,003	-2,058
Jul	101,529	110,073	-8,544	110,101	-8,572
Aug	133,429	128,894	4,535	128,602	4,827
Sep	150,975	145,282	5,693	144,973	6,002
Oct	144,720	155,115	-10,395	155,141	-10,421
Nov	140,880	135,680	5,200	135,858	5,022
Dec	221,502	226,082	-4,580	227,262	-5,760
Total	1,931,981	1,931,982	**-1**	1,933,688	**-1,707**

Table E-4 shows that ordinary regression analysis results in a net determination bias (NDB) of more than the acceptable cut-off criterion of 0.005% given in ASHRAE Guideline 14.[33] The weighted regression provides a NDB that meets this criterion and could be improved by using more precise estimates of the coefficients.

Table E-4. Comparison of NDB between weighted and ordinary regression analysis

Method	NDB
Weighted	-5.8E-07
Ordinary	-8.8E-04

While duplication of observations is a simple method for performing weighted regression, it should be noted that it produces artificially high R^2 values and *t*-statistics for independent variables. In these cases, ordinary regression should be applied for the screening of competing models and the selection of independent variables, with weighted regression applied as a final step to dial in the coefficient values on the selected model (for the purpose of minimizing NDB). However, a true weighted least-squares regression analysis (i.e., one that doesn't depend on an *ordinary* least-squares regression of duplicated data) should properly account for the diagonal matrix, W_{ii}, in its R^2 and *t*-statistic calculations. In such cases, it is better to screen competing models using the weighted regression analysis and statistics.

[33] ASHRAE, Annex B.

Appendix F – Summary of Competing Models

An example summary of competing models is shown below.

Table F-1. Example of competing model summary

No.	Freq.	Period	Days in Baseline Period	R^2	Adj. R^2	CV-RMSE (%)	Auto-corr. Coeff.	FSU (5.0% savings, 80% CL)	Net Det. Bias	Variables	Coefficients	T-value	Comments
1	Daily	9/1/2014 to 8/31/2016	365	0.871	0.865	5.6%	0.280	19.5%	1.08E-14	Constant	37,340	10.3	Linear model with both production variables and temperature.
										Temp	560	7.5	
										Variable 1	1,103	3.7	
										Variable 2	1,200	7.6	
2	Daily	9/1/2014 to 8/31/2016	365	0.882	0.876	5.4%	0.270	18.6%	-1.01E-14	Constant	33,288	9.6	Change point model with boh production variables.
										Temp	1,997	9.9	
										Change-point	53		
										Variable 1	1,003	1.2	
										Variable 2	1,178	8.5	
3	Daily	9/1/2014 to 8/31/2016	365	0.912	0.901	5.1%	0.250	17.5%	3.98E-14	Constant	27,643	6.7	This model includes both a change-point and an interaction term for the two production variables. This model provided the best fit and accounts for the effects of the production lines on each other. **Final Model.**
										Temp	1,875	7.9	
										Change-point	53	2.4	
										Variable 1	978	2.0	
										Variable 2	1,009	7.3	
										Variable 1 x Variable 2	0.045	2.9	

Appendix G – Ranges for Outlier Screening and Variable Limits

This guideline makes use of both ± 3σ and ± 4σ to detect outliers. In general, ± 3σ is used for independent variable screening when a normal distribution of values is expected. When residuals are under review, however, a wider range of ± 4σ is generally considered appropriate.[34] The probability that all residuals will lie within ± 4σ is 99.99%. Residuals that exceed ± 4σ should be reviewed and documented. These are general recommendations, and under certain circumstances other ranges may be applicable.

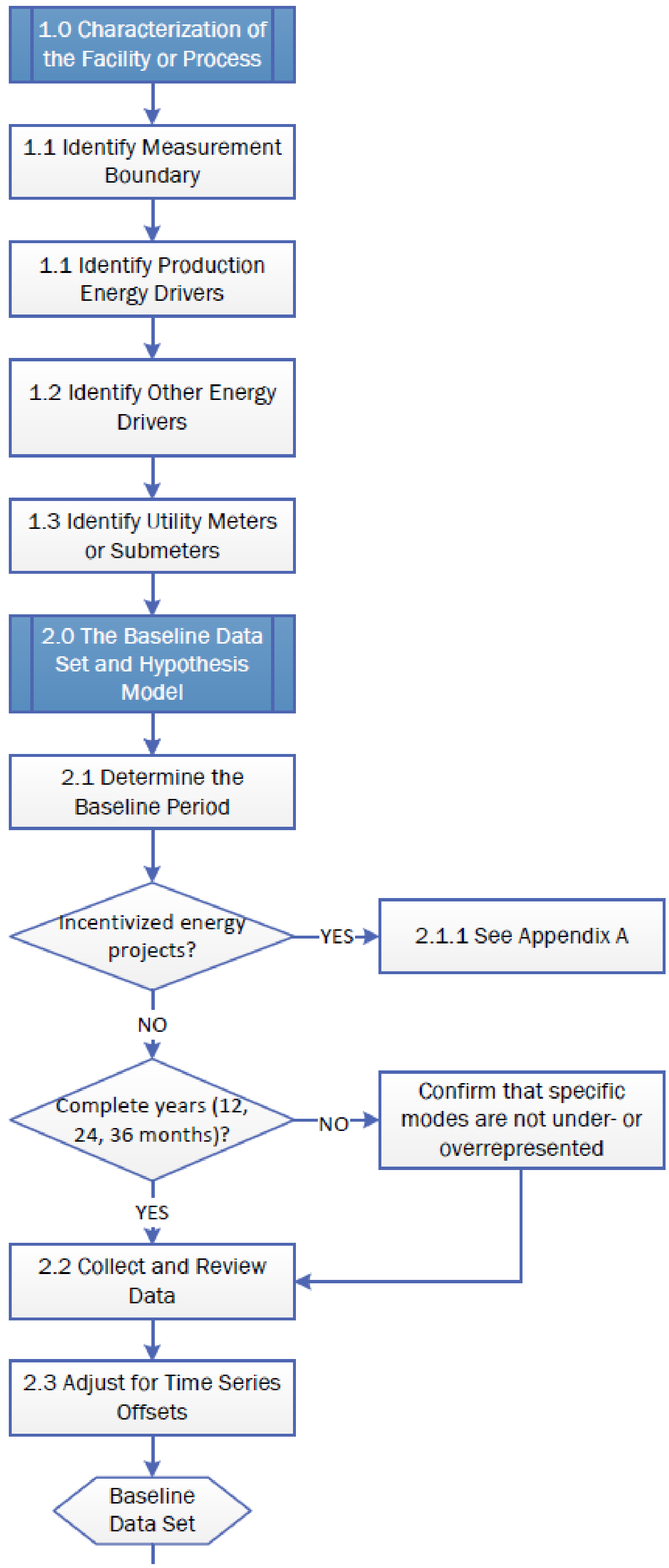
Below is a summary of where ± 3σ and ± 4σ are generally used, and how sigma is calculated.

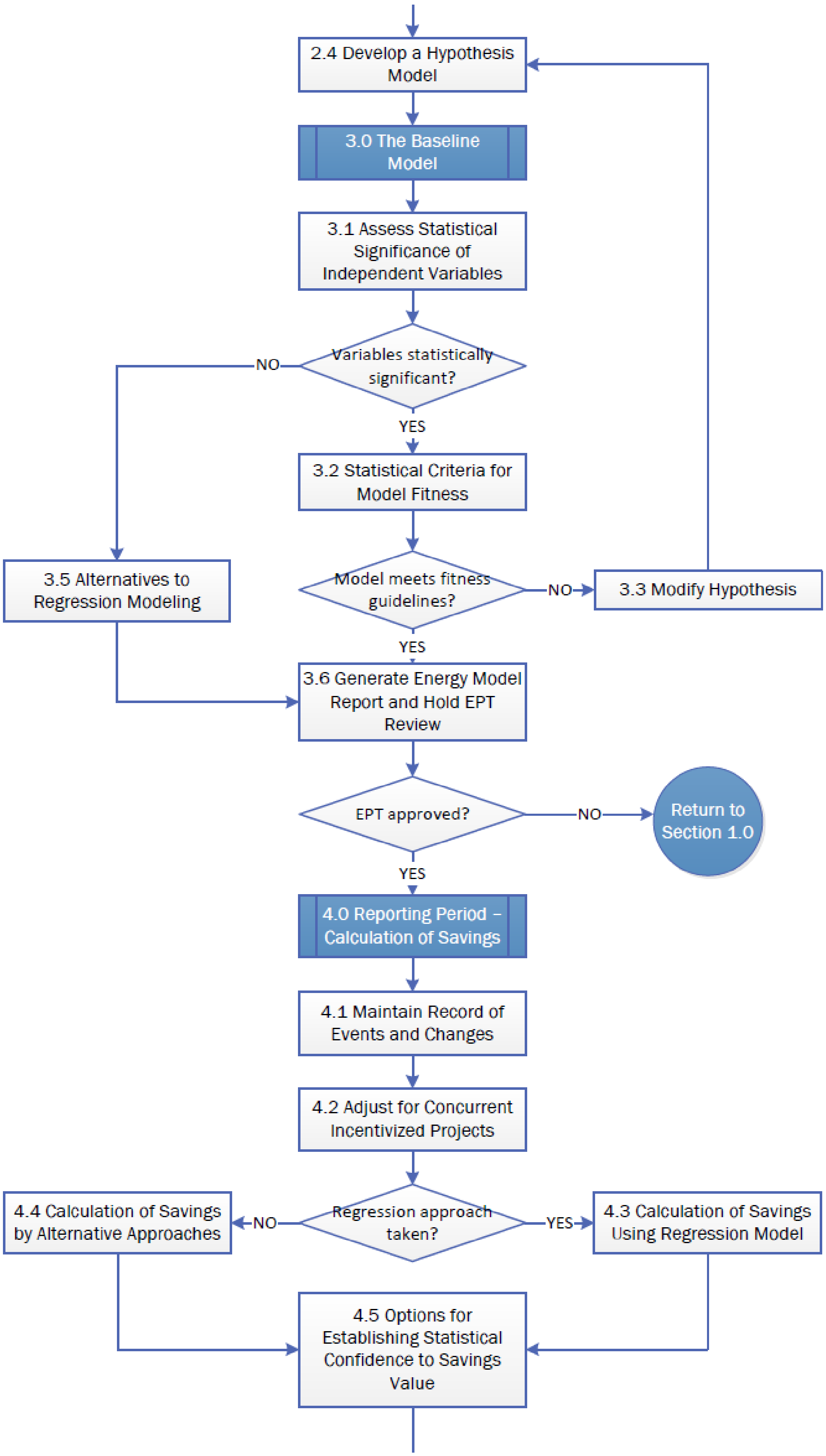
Table G-1. Summary of ± 3σ and ± 4σ

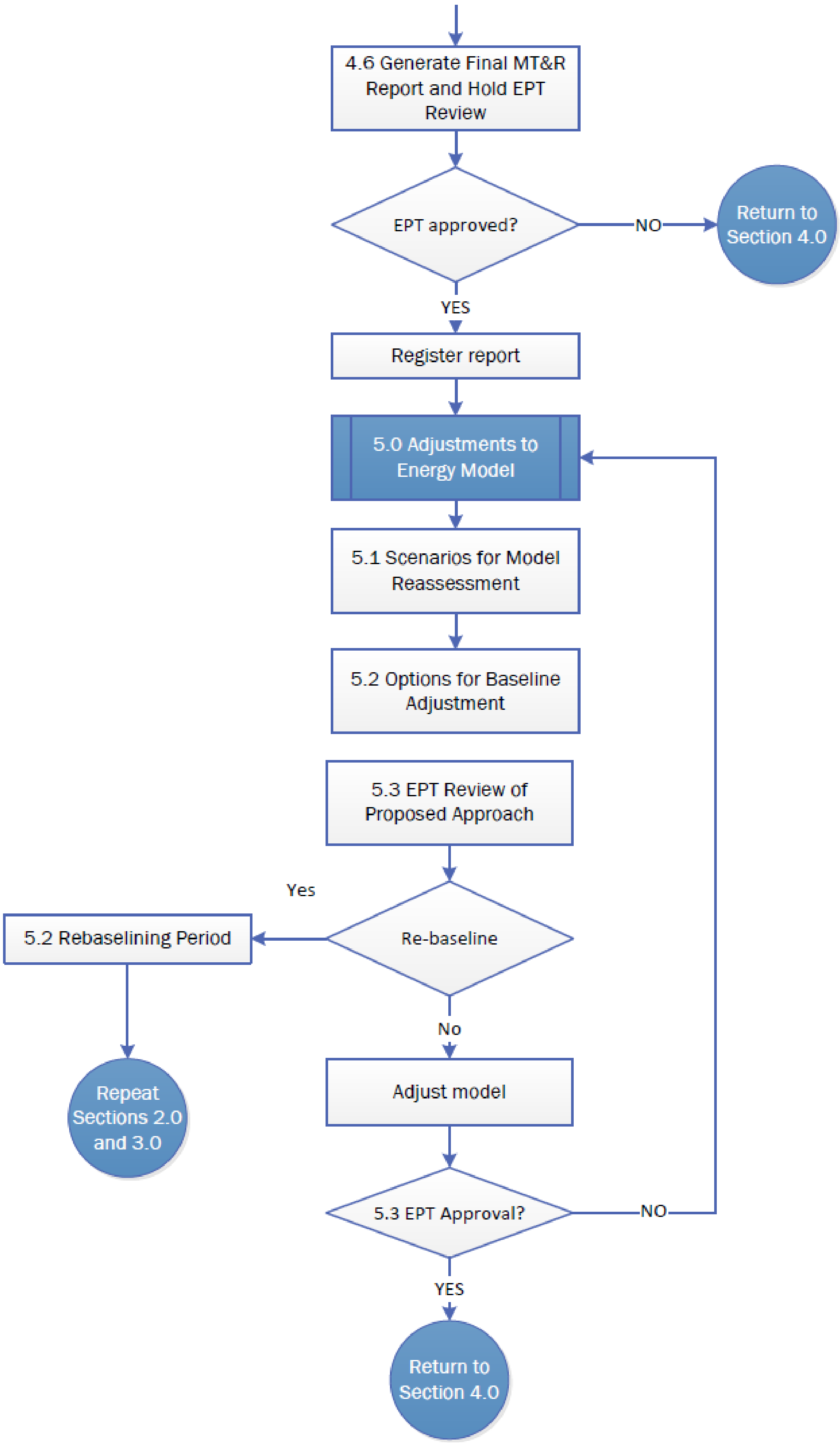
RANGE	APPLICATION	ESTIMATE OF σ
± 3σ	• Set control limits for review of raw energy and energy driver data (Section 2.2). • Validate operational parameter values during savings period when using mean model (Section 3.5.2). • Validate limits of dependent variables for a single mode regression model (Sections 4.3 and 5.1).	Standard deviation of the variable for the baseline period.
± 4σ	• Baseline model residuals outlier detection (Section 3.4)	√MSE
± 4σ	• Performance period outlier detection: Performance period residuals (Section 4.3).	Standard deviation of the energy savings for the respective period.

34 Kutner, Nachtsheim, Neter. 2004. *Applied Linear Regression Models, Fourth Edition.* McGraw Hill. p 108.

Appendix H – MT&R Decision Tree




2.4 Develop a Hypothesis Model
3.0 The Baseline Model
3.1 Assess Statistical Significance of Independent Variables
Variables statistically significant?
NO
YES
3.2 Statistical Criteria for Model Fitness
3.5 Alternatives to Regression Modeling
Model meets fitness guidelines?
NO
3.3 Modify Hypothesis
YES
3.6 Generate Energy Model Report and Hold EPT Review
EPT approved?
NO
Return to Section 1.0
YES
4.0 Reporting Period – Calculation of Savings
4.1 Maintain Record of Events and Changes
4.2 Adjust for Concurrent Incentivized Projects
4.4 Calculation of Savings by Alternative Approaches
NO
Regression approach taken?
YES
4.3 Calculation of Savings Using Regression Model
4.5 Options for Establishing Statistical Confidence to Savings Value


4.6 Generate Final MT&R Report and Hold EPT Review
EPT approved?
NO
Return to Section 4.0
YES
Register report
5.0 Adjustments to Energy Model
5.1 Scenarios for Model Reassessment
5.2 Options for Baseline Adjustment
5.3 EPT Review of Proposed Approach
Yes
Re-baseline
5.2 Rebaselining Period
No
Repeat Sections 2.0 and 3.0
Adjust model
5.3 EPT Approval?
NO
YES
Return to Section 4.0

Appendix I – Revision History

REV	RELEASE DATE	SECTION	CHANGES
1.0	Apr 2010		• New Document
2.0	May 2010		• Addressed feedback from BPA Planning and CADMUS Group (Document Dated April 15, 2010).
3.0	Mar 2012	General	• Incorporated Document Objective, clearly stating ownership by ESI EPT team. • Added various appendixes and illustrations, including Glossary of Terms. • Added revision history.
		Section 1	• Added a requirement that the effect of ambient temperature should always be tested for statistical significance. • Clarified requirement for calibration of in-house submeters that don't match revenue meter boundary.
		Section 2	• Clarified strong preference for including even intervals of annual cycles in baseline period. • Included specific guidelines for adjusting for incentivized or non-incentivized EEMs that were installed during the baseline period. • Added additional guidance and illustration for outlier removal, and time-series adjustments. • Included discussion of change-point models. • Added a discussion of multicollinearity.
		Section 3	• Added a requirement to assess auto-correlation of the residuals. • Added a requirement to calculate Net Determination Bias of the residuals. • Added a requirement to calculate adjusted R^2. • Included specific options for "Alternatives to Regression Modeling."
		Section 4	• Added guidance on adjustments for concurrent incentivized projects during the "reporting period." • Added discussion of model uncertainty.
		Section 5	• Added a section that outlines specific options for baseline adjustment.
4.0	Sep 2013	Section 2.2	• Changed data screening criteria from three standard deviations to four standard deviations. • Changed reference for data screening. • Eliminated graph in Figure 1.
		Section 2.4	• Adding clarifying language for multicollinearity. • Added reference for multicollinearity.
		Section 3.2	• Replaced Figure 6 with new figure. • Added Durbin-Watson test statistic.
		Section 3.4	• Added section.
		Section 3.5.1	• Added section.
		Section 3.5.2	• Terminology change from mean-shift to mean model.
		Section 4.3	• New figure for Figure 8.
		Section 4.5.2	• Added section.
		Section 4.5.3	• Added section.
		Section 6.0	• Added section.

REV	RELEASE DATE	SECTION	CHANGES
5.0	Feb 2015	Section 1.1	• Added content regarding the measurement boundary and accounting for all energy and mass flows crossing the boundary. • Added content about the inclusion of process parameters within the energy mode. • Added content regarding the handling of data from control systems. Included Figure 4 and referenced weighted regression.
		Section 1.2	• Added section.
		Section 2.2	• Added section.
		Section 4.4.3	• Added section: Savings Calculation by Bottom-Up Approach.
		Section 4.4.4	• Added section: Savings Calculation by KPI Based Classification.
		Appendix E	• Added clarifying language about using weighted regression to determine coefficient values.
		Appendix F	• Added Appendix F: KPI Bin Model.
		Appendix G	• Added Appendix G: Summary of Competing Models.
6.0	June 2017	Section 1.2	• Eliminated reference to dialoguing with key contractors.
		Section 1.3	• Require more rigorous documentation when temperature is omitted from model. • Revised Figure 2. • Replaced Washington State University Agricultural Weather Network weather source with Weather Underground.
		Section 2.1	• Added bullet for baseline period for re-enrollment. • Clarification of weather dependent models.
		Section 2.2	• Emphasized collecting and screening of data. • Revised Figure 3.
		Section 2.4	• Replaced figure 6 with a more representative data set. • Added reference to degree day models. • Added reference to exploring non-linear and interactive effects.
		Section 3.4	• Revised Figure 10.
		Section 3.5.1	• Revised application of back-cast method.
		Section 4.3	• Modified default method for establishing valid range to +/- 3 sigma. • Added clarification of how to calculate savings when data is out of range and savings are high.
		Section 4.4.1	• Revised savings calculations for back-cast method.
		Section 4.4.3	• Added section.
		Section 4.4.4	• Revised the use of the bottom-up approach.
		Section 4.5.3	• Revised t-test.
		Section 5.1.1	• Added section.
		Section 5.1.2	• Added section.
		Section 5.1.3	• Added section.
		Section 5.2.1	• Added section.
		Section 5.2.2	• Added section.
		Section 5.2.3	• Added section.

REV	RELEASE DATE	SECTION	CHANGES
		Section 6.4	• Added section.
		Appendix H	• Revised Flow Diagram
7.0	Oct 2018	Introduction	• Added introduction. Six major section titles changed in document to reflect those listed in the introduction.
		Section 1.1	• Revised Figure 1.
		Section 1.3	• Revised Figure 2.
		Section 1.4	• ASHRAE reference updated.
		Section 2.1	• Minimum number of baseline data points is a guideline.
		Section 2.2	• Data screening changed from 4σ to 3σ. • Added permissibility of interpolating or replacing missing data.
		Section 2.4	• Added key factors for considering a degree-day approach. • Added ways to identify and address multi-collinearity.
		Section 2.4.1	• Added section, including Table 2.
		Section 3.2	• Replaced ASHRAE reference for R^2 with SEP reference.
		Section 3.4	• Added clarifying language for applying ±4σ control limits to screening of residuals.
		Section 4.2	• Changed "in-service" date to M&V start date when adjusting for concurrent incentivized projects.
		Section 4.3	• Added clarifying comments as when to use 3σ and 10% of range for valid limits of the model. • Added option to use expected value of savings when an acceptable capping limit cannot be applied.
		Section 4.4.5	• Savings Calculation by Key Performance Indicator (KPI) Bin Model has been removed.
		Section 4.5.3	• Revised application of t-test from one-sided to two-sided test.
		Section 6.0	• Section generalized for SEM projects. • Added fourth bullet point.
		Appendix B	• Changed project M&V end date to M&V start date when adjusting for incentivized projects in the reporting period.
		Appendix D	• Added heating and cooling degree-days.
		Appendix H	• Appendix added.
		KPI Bin Method	• Removed appendix for KPI Bin Method (Formerly Appendix F).